JPRS: 3532

19 July 1960

SOVIET DEVELOPMENTS IN INFORMATION PROCESSING AND MACHINE TRANSLATION

NO 20

U. S. JOINT PUBLICATIONS RESEARCH SERVICE
205 EAST 42nd STREET, SUITE 300
NEW YORK 17, N.Y.

F O R E W A R D

This publication was prepared under contract by the UNITED STATES JOINT PUBLICATIONS RESEARCH SERVICE, a federal government organization established to service the translation and research needs of the various government departments.

JPRS: 3532

CSO: 3901-D/20

SOVIET DEVELOPMENTS IN INFORMATION PROCESSING AND MACHINE TRANSLATION

FOREWORD

This translation series presents information from Soviet literature on developments in the following fields in information processing and machine translation: organization, storage and retrieval of information; coding; programming; character and pattern recogniton; logical design of information and translation machines; linguistic analysis with machine translation application; mathematical and applied linguistics; machine translation studies. The series is published as an aid to U. S. Government research.

Previously issued JPRS reports on this subject include:

JPRS: 68, 241, 319, 355, 379, 387, 487, 621, 646, 662, 705, 729, 863, 893, 925, 991, 992, 1006, 1029, 1130, 1131, 1132, 1133, 3225, 3300, 3356, 3433, 3502.

JPRS: 3532

CSO: 3901-D/20

THE TRANSLATION MACHINE OF P. P. TROYANSKIY

[Following is a translation of a portion (pages 1-40) of a Russian-language booklet comprising a collection of materials on a translation machine proposed by P. P. Troyanskiy in 1933. The booklet, Perevodnaya mashina P. P. Troyanskogo, (The Translation Machine of P. P. Troyanskiy), was edited by D. Yu. Panov and published by the Publishing House of the Academy of Sciences USSR, Moscow, 1959, 51 pages.]

In 1933 P. P. Troyanskiy applied for a patent on his invention of a "Machine to Produce Automatically Finished Printed Translations from One into Several Languages Simultaneously and Requiring Only Literary Editing." He was granted USSR patent No 40995 for a "Machine to Select and Print Words in Translating from One Language into Another," dated 5 September 1933. This was the first proposal on record for the construction of a special translation machine.

A good deal of work has been done in recent years on automatic translation in the USSR and abroad. Hence, some of Troyanskiy's ideas expressed back in 1933 and currently the subject of discussion among scholars are of considerable interest.

Accordingly, the Presidium of the Academy of Sciences USSR set up a committee in 1957 to study the materials made available by Z. N. Smirnova-Troyanskaya, Troyanskiy's widow, and to prepare some of them for publication.

Petr Petrovich Troyanskiy was born in January 1894 into the family of a railroad shop worker in the city of Orenburg. There were 13 other children and life was very hard. Troyanskiy graduated from parochial school and obtained his high school diploma by passing the examinations, although he did not attend classes. He then went on to St. Petersburg University, where he earned a living by tutoring. He was unable to complete his university studies due to the outbreak of World War I. He attended the Institute of the Red Professorate after the October Revolution. He then taught social sciences and the history of science and technology in higher educational institutions. He also did a good deal of writing for the Technical Encyclopedia and the Great Soviet Encyclopedia. During those years he spent increasing amounts of time trying to realize his idea of a translation machine. A serious heart condition prevented him from completing his work on the mechanization of translation, which he regarded as his lifework. He died on 24 May 1950.

The materials on P. P. Troyanskiy's translation machine are published with the permission of Z. N. Smirnova-Troyanskaya and include commentaries and notes by members of the committee of the Presidium of the Academy of Sciences USSR. The text of Troyanskiy's papers is printed unchanged from his manuscripts. I. K. Bel'skaya prepared the linguistic portion of the materials and wrote the commentary on his article dealing with the linguistic principles underlying the design of his translation machine. L. N. Korolev and D. Yu. Panov handled the technical portion and wrote the commentary on the technical materials. D. Yu. Panov served as the general editor.

I. K. Bel'skaya, L. N. Korolev, and D. Yu. Panov

PART I. LINGUISTIC MATERIALS

P. P. SMIRNOV-TROYANSKIY

A TRANSLATION MACHINE BASED ON THE MONOLINGUAL METHOD

1. Logical Analysis

In developing a new method of translation I started with the logical structure common to all languages: everywhere a subject is a subject, a predicate is a predicate, an object is an object, etc. This common logical structure brings languages close together, despite the variety of structures, grammars, and vocabularies, and makes it possible to overcome the differences between them.

On the basis of this unity I created a so-called form of logical analysis common to all languages as a kind of intermediate text in the translation process.

The operation of obtaining a form of logical analysis requires that all the words be used in their original grammatical forms, i.e., in the nominative case, singular number, for all declinables and in the infinitive for all conjugable verbs, when transforming a natural text, e.g., text A, into a form of logical analysis, e.g., into text A^1. At the same time, the so-called invariants, i.e., the nonconjugable and nondeclinable words (members) of the sentence such as prepositions, conjunctions, adverbs, and interjections, retain their one inherent grammatical form. However, the task of imparting meaning to the disconnected words, i.e., the work of collecting groups of words into logical units and connected clauses, is performed by the symbols of logical analysis that accompany each declinable or conjugable word in the sentence and sometimes even invariant words (that is, when they function as the principal members of a sentence, e.g., "Hurrah burst out in the distance," [daleche gryanula ura], where "hurrah" is the subject and in similar rare cases).

There are some 25 international symbols of logical analysis common to all languages. They are used in some 100 different combinations and are capable of conveying all relations without exception and the subtlest shades of human thought. Expressing thought with words and concepts, they can ensure completely accurate translation into other languages without any distortion of meaning.

The symbols of logical analysis show the logical syntactic significance (the semantic role) of a word in a sentence and at the same time they free the sentence from its mandatory word order, which, as it has turned out, is well within their capabilities and very convenient for the new translation technology.

In making a logical analysis I, so to speak, take out of parentheses that wherein languages resemble one another and leave within parentheses that wherein they differ from one another. The transformed text is a text in the form of a logical analysis. The machine operates on the text thus transformed. (Cf. Part II for Troyanskiy's machine.)

That which is taken out of parentheses (i.e., the same representation of the semantic, logical connection between words in a sentence that is common to all languages and between members of a clause -- a connection that is likewise identical and common to all languages) is shown and expressed by the symbols of logical analysis. On the other hand, that which is left within parentheses (that wherein languages differ from one another) is the vocabulary in its original grammatical form (i.e., in the nominative case, singular number for declinable words, in the infinitive for conjugable verbs, and in their single inherent form for invariants). Many simple and convenient procedures for all kinds of translation are included in the special instructions for performing logical analysis.

It is apparent that the form of logical analysis (form A^1 or form B^1, C^1, ... K^1 in other languages) is extremely rich, for it creates many unexpected possibilities for extensive simplification of the entire technological process of translation and for transforming it into something that can be used by everyone.

Moreover, utilization of logical analysis gave rise to 12 completely new types of technological processes for translation that were unknown or rather were completely lacking hitherto. Each of these solves a certain practical problem and enables a person to use several foreign languages even though he knows only his native language.

Any mechanization of labor processes introduces its own laws. It makes particular demands on the material. When using a machine (this is true of all machines without exception), the material to be processed must first be suitably treated either for size or for chemical or some other special technical properties or conditions. Should there be some gross technological disparity between machine and material, the machine will refuse to work.

Logical analysis is a similar kind of preliminary operation on the material to be processed by the translation machine. It is an integral part of machine translation technology. Like any other ordinary machine, the translation machine likewise demands of people desiring to obtain normal results from it that they observe its rules, specifically that they supply for processing not any raw material, but only special material which has first been suitably adapted. A text in the form of a logical analysis is the material made ready for processing by the translation machine.

Thus, like any ordinary machine, the translation machine will operate normally and produce the desired product when it is duly supplied with material ready for processing, i.e., a logical analysis.

2. The Technological Process of Translation and Its Three Operations

This is how the process is carried out in practice.

In reorganizing the technology I divided the translation and editing processes between two persons, one of whom knows only the language of the original and the other only the language of the translation, i.e., between two persons neither of whom knows both languages.

I then used the conventional system of symbols of logical analysis and divided the translation process into three separate operations.

In the first or monolingual operation -- the replacement of A with A^1 -- the original text A to be translated is replaced with text A^1 in the same language, but in the form of a logical analysis. This initial operation $A-A^1$ is performed by the first translator who knows only his native language, here the language of the original.

In the second or bilingual operation -- the replacement of A^1 with B^1 -- text A^1 in the form of a logical analysis in the language of the original is replaced with text B^1, which is also in the form of a logical analysis, but in the language of the translation. This operation A^1-B^1 is performed by the machine. There is no need here of any logical work on the text. The original forms of one language are mechanically exchanged for the original forms of another language, but the symbols of logical analysis are converted accordingly, likewise in mechanical fashion, into analogous forms of the other language. The result is an exchange of texts in the form of a logical analysis, A^1-B^1.

The third operation, like the first, is monolingual and involves the replacement of B^1 with B. It features the second translator whose native language is the language of the translation and who, like the first translator, knows no other language. The form of logical analysis B^1 produced by the machine is replaced by the second translator with the grammatical form B in his own language, the language of the translation. The second translator, therefore, performs operation B^1-B, the reverse of operation $A-A^1$ performed by the first translator on his own text. However, the person who performs the replacement operation in one direction, i.e., $A-A^1$ (or $B-B^1$), can also perform the reverse

operation, i.e., A^1-A (or B^1-B). The second translator is just like the first one when the original text is translated from his native language into other languages.

Thus, two translators, each of whom knows only one language, his mother tongue, together execute the entire translation from beginning to end.

Despite this change in the procedure and nature of the translation process (use of the form and symbols of logical analysis, division of the work between two monolingual translators, possibility -- cf. below -- of using several incomplete translation cycles, etc.), both my monolingual translation process and my monolingual editing operation are performed normally and the results are the same as those obtained by the old bilingual method where each translator must know two languages -- the original and the one into which the translation is to be made. The main differences are that the translation costs are enormously reduced, the scope of translation is maximally enlarged due to its general accessibility, and the actual translation process is mechanized and automatized on the basis of a new monolingual technique.

The technological process is so organized that the second operation A^1-B^1 (or the reverse operation B^1-A^1) can be executed on the machine by either the first or second translator, or by an absolute stranger who knows neither the language of the original nor the language of the translation; for, if there is a logical analysis, the operation is quite simple.

In recent versions of a complete automaton, where the work of the first translator is done directly on punched tape, the translation machine is controlled by the tape without human intervention.

The translation process therefore consists of four successive stages A-A^1-B^1-B, or in the reverse direction, B-B^1-A^1-A.

However, the machine can do more than that.

The first translator can translate texts on the machine not only into any one language but also into several languages simultaneously in a series of consecutive stages beginning with the first operation A-A^1, then -- on the machine -- the second operation A^1-(B^1 + C^1 + ... + K^1). This second operation of multilingual translation on the machine can also be performed by the second translator, if he has in his hands the text in the form of A^1 prepared by the first translator, or even by a total stranger who knows neither the language of the original nor the language of the translation as long as he has the text of A^1. In this second operation of multilingual translation the machine itself makes the necessary exchange of texts: it replaces the text of A^1 in the basic grammatical forms of the original with texts in the languages of the translation, also in the basic forms, and automatically supplies them with the symbols of logical analysis in accordance with the form of A^1.

Finally, the use of several second translators (their number: B + C... + K = T) makes it possible to carry out the third operation of multilingual translation, i.e., (B^1 + C^1 + ... + K^1) - (B + C + ... K),

after which the text emerges in the correct grammatical forms of the various languages.

3. The Simplicity of Performing the Operation of Logical Analysis

A monolingual translator will have no difficulty in making a logical analysis (the A-A^1 type) or the reverse process, i.e., transforming a logical analysis into the grammatical forms of some language (i.e., the A^1-A type of operation). This is rather like the elementary school exercises in one's native language called parsing a text "by the parts of speech" and "by the parts of a sentence." The only difference is that in parsing the results are recorded more clumsily than those of a logical analysis designed for translation purposes.

Here is an example of parsing: Ptichka letayet. (the bird flies). Ptichka is a noun, feminine, nominative case, singular number, answering the question kto (who) = subject. Letayet is a verb, present tense, singular number, third person, indicative mood, answering the question chto delayet (what does it do) = predicate.

In the same example my recording of the logical analysis of Ptichka letayet is shorter and simpler. Everything superfluous is discarded, but the full idea is conveyed, i.e., Ptichka - o, letat' - a, which is all that is needed to record the A-A^1 type of operation. If the kind of parsing described above is easy for school children, the recording of my logical analysis is even easier, especially for a literate adult. The same holds true for the reverse operation, i.e., the A^1-A type.

4. Some New Features

As I was working out my ideas on monolingual translation, increasing numbers of new features gradually began to appear and accumulate, thus testifying to their soundness, for unsound ideas are incapable of development.

1. I began with the full possibility (described above) of two persons each knowing only one language and producing a normal translation after going through the complete cycle A-A^1-B^1-B (or through the reverse cycle B-B^1-A^1-A, if the second was the original language). The results were in no way distinguishable from high quality translations produced hitherto by the old bilingual method, i.e., by persons knowing the two languages involved.

2. I then found it completely satisfactory to publish texts only in the form of a logical analysis, irrespective, of course, of the method of obtaining these texts: either as a result of the work of the first translator after operation A-A^1 (when, for example, editors supply the symbols of logical analysis in footnotes to scientific books and articles, thereby issuing normal texts along with their logical analysis (see samples of translations below) so as to make them ready for translation into other languages) or as a result of machine operation

A^1-B^1 (or B^1-A^1) for use, let us say, as telegraphic text that is transmitted to another country in the language of the sending country, but is received by the addressee in the language of his own country, etc.

3. It will be possible for two persons in different countries to correspond. The sender will be able to write in the language of the receiver without knowing his language (by going through the A-A^1-B^1 type translation procedure and sending out text B^1) while the receiver answers in the sender's language with which he too is unfamiliar (carrying out the B-B^1-A^1 procedure and replying with text A^1).

4. Given articles and books published with the form of logical analysis, one can translate them without knowing the language of the original by using procedure A^1-B^1-B or B^1-A^1-A, etc.

5. Translation into several languages simultaneously is possible by A-A^1-(B^1+C^1+...K^1) - (B+C+...+K).

6. It has been found that translation costs can be reduced to as little as one percent of the previous figure. Obviously, then, the greater the number of languages translated by machine, the cheaper will be the translation in each individual language.

7. It is now possible for any writer, scholar, or journalist to translate his own works into a foreign language or even into several languages at the same time by using A-A^1-B^1 or A-A^1-(B^1+C^1+...K^1) and issuing the materials in the form of a logical analysis, i.e., giving the second translators 90 percent of the finished translation and leaving it to them to carry out only procedure (B^1+C^1+...K^1) - (B+C+...+K), or just the remaining 10 percent of the work -- the operation involved in changing the texts given in the form of a logical analysis into their full grammatical forms.

8. It will improve the quality of the editorial work at all three stages A-A^1, A^1-B^1, and B^1-B, which is now done only by monolingual editors.

In case of doubt as to the correct transmission of some part of a translation, each monolingual editor (whose function may generally be performed by the same monolingual translator) can clear up the confusion in the translation by corresponding with his opposite number in the way described in (3) above, i.e., using A-A^1-B^1 and B-B^1-A at either end of the translation operation.

The accuracy and quality of operation A^1-B^1, which is performed by the machine without human intervention, is guaranteed -- from the editorial point of view -- by the skillful work of linguists on the machine dictionary. Specialists in applied linguistics watch to see that the vocabulary incorporated in the machine is in line with the modern development of the languages. Any necessary changes in the dictionary can be made as easily as changing a fuse in a fuse box.

9. Translation is brought within general reach of the masses. Any literate person can now translate.

Important Note: My monolingual translation method does not do away with the need of a specialized education in order to understand a specialized text in one's native language. Oddly enough, this

has been offered as a criticism of my method and translation machine. The fact is that the peculiarities of any specialized field have to be learned. This is true even when the bilingual method is used and no one, least of all myself, says the method is at fault.

10. Finally, mechanization and automation of translation is now a tested reality in the form of a model that functions automatically.

For the design and operational features of this machine, see the technical description in Part II.

5. Table Comparing the Monolingual and Bilingual Translation Methods

I have compiled Table 1 to make clear the significance of my work and to show graphically my innovations as compared with the translation methods used hitherto. It presents 15 formulas representing the various procedures involved. They are grouped by the form of text with which the translation process begins and by the form of text with which it ends.

The table also shows the way the translation processes differ in terms of the number of translators required for each and whether each man knows one or two languages.

The formulas indicate the successive forms that the text goes through in translation. For example, No 10, $A-A^1-B^1-B$, means that an actual text A is first replaced by a text A^1 in the same language but in the form of a logical analysis. The resultant text A is replaced by an equivalent text B^1, also in the form of a logical analysis but in the other language. And, finally, this text in the form of B^1 is replaced by text B, i.e., a text with all the normal grammatical forms of language B. The remaining formulas present the same transformation of text.

The table clearly contrasts both methods, showing the distinguishing features that make it possible to evaluate them. Some other characteristics are shown in Table 2.

TABLE 1

Comparison of Monolingual and Bilingual Translation Methods

Groups of processes	Order of processes	Form with which the translation process begins	Kinds of translation processes (formulas of translation processes)	Form with which the translation process ends	No of translators required	No of languages each translator knows
			MONOLINGUAL METHOD			
1	1	FLA	$A^1 - B^1$	FLA	0	0
	2		$B^1 - A^1$		0	0
	3		$A^1 - (B^1 + C^1 + \ldots + K^1)$		0	0
2	4	NGF	$A^1 - B^1 - C$	NGF	1	1
	5		$B^1 - A^1 - A$		1	1
	6		$A^1 - (B^1 + C^1 + \ldots + K^1)$ $- (B + C + \ldots + K)$		T	1
3	7	NGF	$A - A^1 - B^1$	FLA	1	1
	8		$B - B^1 - A^1$		1	1
	9		$A - A^1 - (B^1 + C^1 + \ldots + K^1)$		1	1
4	10	NGF	$A - A^1 - B^1 - B$	NGF	2	1
	11		$B - B^1 - A^1 - A$		2	1
	12		$A - A^1 - (B^1 + C^1 + \ldots + K^1)$ $- (B + C \ldots + K)$		T+1	1
			BILINGUAL METHOD			
5	13	NGF	$A - B$	NGF	1	2
	14		$B - A$		1	2
	15		$A - (B + C + \ldots + K)$		T	2

Explanation of the Symbols in Comparative Table 1

1. FLA = form of logical analysis
   NGF = native grammatical form
2. A, B, C,...K = texts of various languages in their native grammatical form.
3. $A^1, B^1, C^1 \ldots K^1$ = texts in the form of logical analysis.

4. For the types of translation process under 1, 2, and 3, the number of translators and number of languages = 0. This means that these processes can be completely automatized, i.e., no human is required; under nonautomatized conditions the translators need not know either the language of the original or the language into which the translation is to be made, i.e., none (thus the 0) of the languages involved in these processes.

5. In the formulas of the processes in parentheses (nos. 3, 6, 9, 12, and 15) T = the number of languages into which the translation is made, T+1 = the number of languages + the language of the original.

Contents of Table 1

The table shows that among the 12 monolingual translation processes three (10, 11, and 12) are comparable to three bilingual processes (13, 14, 15) i.e., they produce identical results but by different methods. Moreover, the comparison does not favor the old, bilingual method.

Comparison 1 -- bilingual process 13, A-B, which requires one man knowing two languages. The same result is obtained with the monolingual method using process 10, $A-A^1-B^1-B$. Two persons are required, but each needs to know only one language.

Comparison 2 -- bilingual process 15, A-(B+C+...K) with translation into T languages. T persons are required, each knowing two languages. On the other hand, monolingual process 12, $A-A^1-(B^1+C^1+...K^1) - (B+C+...+K)$, with translation likewise into T languages, requires T+1 persons, but each knows only one language. The advantage clearly rests with the monolingual method.

Comparison 3 -- bilingual process 14, B-A, which proceeds in the opposite direction from the similar bilingual process 13. The process B-A requires one person knowing two languages. Each bilingual process is characterized by the fact that a person who can easily translate from language A into language B according to formula A-B usually finds it very difficult, and often he even flatly refuses, to carry out the reverse process of translating from language B into language A according to formula B-A. Consequently, the A-B and B-A processes in the bilingual method are not at all equivalent. The results of bilingual process 14, B-A, are comparable to those obtained with monolingual process 11, $B-B^1-A^1-A$. This is likewise a reverse process, i.e., the reverse of monolingual process 10, $A-A^1-B^1-B$. However, with the monolingual method there are no difficulties in carrying out any reverse process. In this case $A-A^1-B^1-B$ and $B-B^1-A^1-A$ are equally simple. Again a clear advantage for the monolingual method.

None of the other translation processes in the monolingual method has any counterpart in the bilingual method. The results of processes 1, 2, 3, 4, 5, 6, 7, 8, and 9 are simply unattainable by the bilingual method with its limited procedures (cf. Table 2).

Translators using the old, bilingual method will not be indifferent to the new, monolingual system because it opens up broad new possibilities for them too. It enables them to translate into a number of languages with which they are completely unfamiliar, something they cannot do by the old method. The monolingual system with all its 12 procedures for the first time makes it possible for people to translate languages they do not know.

6. The most Important Thing in the Comparison

The most important thing in comparing the two methods is perhaps the following. The monolingual method is capable of ensuring the execution of any of the translation processes effected by the bilingual method, that is, processes 13, 14, and 15. In other words, the monolingual method can assume the entire task of the bilingual method. Naturally it does not aim at this, nor does it want to, but clearly it is now feasible to divide the sphere of work between the two methods.

It is quite obvious that the bilingual method cannot by itself carry out any of the 12 translation processes of the monolingual method, that is, it cannot be substituted for the monolingual method nor take the place of any of the new, unusual, necessary, important, practicable operations introduced by the newly created monolingual method (cf. Table 2).

However, I won't argue. Let the old bilingual translators translate without the machine into the languages that they know. But there are languages that they do not know. Before those languages they halt as though faced by a stone wall. This is a dead end for them. They can do nothing with unfamiliar languages.

Now there is a way out -- the monolingual method. By adopting this method they simultaneously become monolingual translators and avoid the dead end.

Besides, 99 percent of the world's literate people know only their mother tongue and do not know foreign languages well enough to be translators. For these people the monolingual method opens up a way to make wide use of foreign languages without knowing them.

This will not prevent them from studying foreign languages, especially those who may have to get along without a machine. The study of foreign languages will continue, and anyone who wishes will be permitted to study.

However, there is an enormous volume of translation work that will have to be handled by the machine. First of all, it is simply impossible to keep up with the increasing flow of material. Why be stubborn when the machine will be able to translate them all into several languages at the same time and promptly put them out in printed form? The translators and editors, you see, will only have to concern themselves with style and polish, for they will be relieved of the mass of hard, routine work of translation.

There is no need to be stubborn. We have master craftsmen who can produce the finest watches with only hand tools. But you would laugh at them if they denied the necessity of using machines to manufacture watches.

It is possible to sow seeds by hand and do it fairly well. But we have mechanical seeders drawn by tractors.

It is also worth noting that the initial models of machines were never perfect. Quite the contrary. Improvements came about gradually — in design, increased size, faster and better operation. It is well to remember the first telephones, typewriters, radios, airplanes, locomotives, and many other machines and instruments. The same thing will happen to the translation machine. It will be helped by the principles underlying linguistic and technological development.

There is and always will be, of course, work which cannot easily be mechanized. For example, surgery on a living human being will scarcely be performed in the near future by a machine. But this does not apply to translation, which is more like computation work that has been mechanized for some time now. They have long since stopped quarreling with machines in this field, although manual labor is still used extensively. Let's not quarrel with the translating machine. It has a job to do.

7. Typical Groups of Translations

Tables 1 and 2 show that there are four typical groups of translation procedures based on the monolingual method.

First group — translation procedures 1, 2, and 3: These are completely mechanical translations that can be produced by a person who is unfamiliar with the languages involved. The result is texts in the form of logical analysis. The translation can be made into one or more languages simultaneously, but if the process is not completely automatized, only one person will be required. The texts obtained by using the first group of procedures may serve as the starting forms for the second group of translation procedures (cf. Tables 1 and 2). The starting forms for the first group of procedures, i.e., procedures 1, 2, and 3, may be either the products of the third group of procedures (cf. Tables 1 and 2), the results of a logical analysis of the A-A^1 or B-B^1 types, or of similar operations performed by monolingual translators while preparing the original text for translation. The texts emerging from procedures 1, 2, and 3 may be used as they are without reduction to the native grammatical forms.

Second group — translation procedures 4, 5, and 6: These are translations which were originally in the forms of logical analysis. It is in these forms that there may be printed, for example, articles and books provided by the editors with the symbols of logical analysis in footnotes and translated by one monolingual translator into his mother tongue with all the native grammatical forms inserted. (For an example of undistorted native text, see Section 14 below.) T

monolingual translators are involved here in the process of making a translation into seven languages simultaneously, which is quite possible, i.e., producing texts entirely in the native grammatical forms. The starting texts for the translation procedure here are the texts produced by the first and third groups of procedures (cf. Tables 1 and 2) and the results of the $A-A^1$ or $B-B^1$ types of logical analysis, etc., which are operations performed by monolingual translators while preparing the original text for translation.

Third group -- translation procedures 7, 8, and 9: These require one monolingual translator to produce from material in his mother tongue texts in the form of a logical analysis in one or more languages unfamiliar to him. The texts resulting from the third group of procedures may serve as starting materials for the first and second groups. However, these texts may also be used just as they emerge from procedures 7, 8, and 9 without all the native grammatical forms being inserted.

Fourth group -- translation procedures 10, 11, and 12: These are complete translations made from the text in the original language forms into one or more languages also in their regular grammatical forms. Translation into one language requires two monolingual persons; translation into T languages requires T+1 monolingual persons.

These conclude the translation procedures in my new, monolingual translation method shown in the comparative table.

Fifth group -- translation procedures 13, 14, and 15: These are based not on the monolingual but on the bilingual method. Translation into one language requires one person with a knowledge of two languages. Translation into T languages requires T translators each of whom is familiar with two languages.

8. Richness of the Monolingual Method

The soundness of the original idea underlying the monolingual method of translation, i.e., the use of logical analysis as an integral part of the process, is shown by the existence of a number of variants and combinations of translation formulas reflecting changes in the actual procedures and the way their work is related to one another. The practicability and value of the original idea is proved by the free "play" of the formulas at the various stages of the operation and by the fact that it enables a great many persons who are literate for the most part only in their own language to do translation. An unsound idea could not have such diverse practical use, which is the test of any new idea.

In time people will become increasingly able to handle a text in the form of a logical analysis. It is not too difficult even now to perform in one's own language an operation of the $A-A^1$ type or the reverse A^1-A operation or to understand A^1 without preliminary processing, i.e., without its being reduced to the forms of A (this can be checked in the examples given below). It will then be possible to make extensive use of an incomplete translation cycle in the form of a

procedure wholly coinciding with the second (purely machine) two-part operation proceeding according to formula 1, A^1-B^1, where the result is a text in the form of B^1 and the starting material is a text in the form of A^1 (or the reverse, which proceeds according to formula 2, B^1-A^1). These procedures can be carried out on the machine by one person who knows neither the language of the original nor the language of the translation.

Moreover, as the ability to use logical analysis becomes more widespread, two expanded but likewise incomplete modifications of the translation cycle will be employed, with operation A^1-B^1 or the reverse operation B^1-A^1 being entirely included within them (note that without A^1-B^1 or the equivalent operation B^1-A^1 translation is generally impossible; I have shown elsewhere that operation A^1-B^1 (or B^1-A^1) occurs in latent form even with the old translation method). These cycles are: first, an incomplete three-part cycle of the A-A^1-B^1 type proceeding according to formula 7, executed by a single person translating from his mother tongue A and producing text B^1, i.e., a text in language B unknown to him in the form of a logical analysis; second, an incomplete three-part cycle of the A^1-B^1-B type proceeding according to formula 4, executed solely by the second translator who works from an unfamiliar language — the text being previously given in this language in the form of logical analysis A^1 — and produces text B in his mother tongue, i.e., with all the regular grammatical forms.

By dividing up and separating in time and space the work of the first translator handling the original text from that of the second translator working on the text obtained from the first one, i.e., using the incomplete three-part cycles mentioned above, we naturally make it possible for communications to be exchanged in foreign languages where neither the sender nor the receiver knows the language of the other. I write in his unfamiliar (to me) language in the form of logical analysis B^1 after first performing the operation of the incomplete cycle A-A^1-B^1 according to formula 7, and receive in return a letter (which he has first handled according to formula 8, B-B^1-A^1) in my language (which he does not know) in the form of logical analysis A^1. This is how a correspondence can be carried on by two people ignorant of each other's language.

The translation procedures indicated in the formulas mentioned above and reflected in the formulas not cited here (cf. Table 1) are reversible, i.e., they may be "read" from either end. Accordingly, each translator can perform both incomplete three-part operations, starting and stopping the procedure in different ways, i.e., in the case of the one from his native language into a foreign language, and vice versa in the case of the other. For the first translator the two procedures look this way: A-A^1-B^1 (according to formula 7) and B^1-A^1-A (according to formula 5). He knows only language A. For the second translator they look exactly the same: B-B^1-A^1 (according to formula 8) and A^1-B^1-B (according to formula 4). He knows only language B.

This too should be noted. Whereas the complete four-part cycle proceeding according to formula 10, $A \rightarrow A^1 \rightarrow B^1 \rightarrow B$ (or the reverse process proceeding according to formula 11, $B \rightarrow B^1 \rightarrow A^1 \rightarrow A$), requires two translators, any of the three incomplete cycles $A^1 \rightarrow B^1$ proceeding according to formula 1, $A^1 \rightarrow B^1 \rightarrow B$ proceeding according to formula 4, and $A \rightarrow A^1 \rightarrow B^1$ proceeding according to formula 7 (or the reverse cycles $B^1 \rightarrow A^1$, $B^1 \rightarrow A^1 \rightarrow A$, and $B \rightarrow B^1 \rightarrow A^1$ proceeding according to formulas 2, 5, and 8, respectively) requires only one translator. Moreover, in the three-part operations $A^1 \rightarrow B^1 \rightarrow B$ and $A \rightarrow A^1 \rightarrow B^1$ (or the reverse $B^1 \rightarrow A^1 \rightarrow A$ and $B \rightarrow B^1 \rightarrow A^1$ proceeding according to formulas 4 and 7, 5 and 8) the translator needs to know only his mother tongue.

If formula 6, $A^1 \rightarrow (B^1 + C^1 + \ldots + K^1) \rightarrow (B + C + \ldots + K)$, or formula 9, $A \rightarrow A^1 \rightarrow (B^1 + C^1 + \ldots + K^1)$, is used, several languages can be translated simultaneously with one translator working in the case of formula 9 and T translators working in the case of formula 6.

Finally, for the shortest translation cycle (of the purely machine type), i.e., in the case of the two-part operation proceeding according to formula 2, $B^1 \rightarrow A^1$, the translation into one language (and where formula 3, $A^1 \rightarrow (B^1 + C^1 + \ldots + K^1)$ is used, into several languages at the same time) can be done by a person unfamiliar with any of the languages involved.

9. Use of Synonyms, Idioms, and Homonyms

The machine supplies from its own dictionary and prints whole groups of synonyms so that the editor has only to cross out whatever is not needed.

Idioms are either "extinguished," i.e., they are replaced with unidiomatic expressions, or left untranslated and then cited unchanged in the foreign language, or a suitable idiom is selected in the language into which the translation is being made.

The correct use of homonyms is assured by the fact that they are not set out in an undifferentiated mass in the manipulatory fields of the machine, but are supplied with a meaning indicator in parentheses, thus:

In the manipulatory field	In the dictionary field
Russian	German
osvobodit^3 (dat^1 svobodu) [to free]	freigeben
osvobodit^2 (otpustit^1) [to let go]	freilassen
osvobodit^1 (izbavit^1) [to release]	erlösen
osvobodit^2 (ot tyazhesti) [to relieve of burden]	entlasten
osvobodit^3 (ot slovo, ot obeshchaniya) [to release from word, promise]	entbinden
osvobodit^2 (ochistit^1) [to clear away]	raumen
osvobodit^1 (mesto, etc.) [to vacate]	freimachen

koso (devich'ya) [girl's braid]	der Zopf
kosa (dlya kos'by) [scythe]	die Sense
kosa (peschanaya) [sand bar]	die Nehrung

TABLE 2

Procedures Carried Out by Monolingual and Bilingual Translators
Using Only Technological Methods
(Yes means they carry out the procedures, no means they do not)

Translators	Translation Proceedures						
	A^1-B^1	B^1-A^1	$A^1-(B^1+C^1+\ldots+K^1)$	A^1-B^1 B^1-A^1 C	B^1-A^1 A	$A^1(+B^1+C^1+\ldots+K^1)-(B+C+\ldots K^1)$	$A-A^1-B^1$
	Group Number						
	1		2				
	Formula Number						
	1	2	3	4	5	6	7
	Monolingual Method						
(a) Requires one person who knows neither the language of the original nor the language of the translation	yes	yes	yes	no	no	no	no
(b) Requires one person who knows only the language of the original	yes	yes	yes	no	yes	no	yes
(c) Requires one person who knows only the language into which the translation is to be made	yes	yes	yes	yes	no	no	no
(d) Requires two persons, one of whom knows the language of the original, the other the language into which the translation is to be made	yes	yes	yes	yes	yes	no	yes

TABLE 2 (continued)

(e) The translation is to be made into T languages; T+1 persons are required, one of whom knows the language of the original, the other ones each of the languages into which the translation is to be made	yes	yes	yes	yes	yes	yes	yes
				Bilingual Method			
(a) Requires one person who knows two languages: the language of the original and the language of the translation	no	no	no	no	no	no	no
(b) The translation is to be made into T languages; T persons are required, each of whom knows two languages: the language of the original and one each of the languages into which the translation is to be made	no	no	no	no	no	no	no

[Lateral continuation of Table 2]

Translation Proceedures

$B \rightarrow B^1 \rightarrow A^1$	$A \rightarrow A^1 \rightarrow (B^1 + C^1 + \ldots + K^1)$	$A \rightarrow A^1 \rightarrow B^1 \rightarrow B$	$B \rightarrow C^1 \rightarrow A^1 \rightarrow A$	$A \rightarrow A^1 \rightarrow (B^1 + C^1 + \ldots + K^1) - (B + C + \ldots + K)$	$A \rightarrow B$	$B \rightarrow A$	$A \rightarrow (B + C + \ldots + K)$

Group Number

3			4				

Formula Number

8	9	10	11	12	13	14	15

Monolingual Method

no	no	no	no	no	no	no	no
no	yes	no	no	no	no	no	no
yes	yes[1]	no	no	no	no	no	no
yes	yes	yes	yes	no	yes[2]	yes[3]	no
yes	yes	yes	yes	yes	yes[2]	yes[3]	yes[4]

Bilingual Method

no	no	no	no	no	yes[5]	no	no
no	no	no	no	no	yes	yes	yes

[1]Using the language of the translation as the original, i.e., with slightly changed formula 9, $B \rightarrow B^1 \rightarrow (A^1 + B^1 \ldots + K^1)$.
[2]By means of the operation proceeding according to formula 10.
[3]By means of the operation proceeding according to formula 11.
[4]By means of the operation proceeding according to formula 12.
[5]As a usual thing. Exceptions are very rare.

An automatized dictionary or a translation machine? The machine that I have proposed, which is based on the monolingual method, is rightly entitled to be called a translation machine because of the three operations that constitute its technological process, i.e., $A \rightarrow A^1$, $A^1 \rightarrow B^1$, and $B^1 \rightarrow B$, the second is a purely translation process. It

embraces different languages and alone produces the actual translation, whereas the first ($A-A^1$) and third (B^1-B) operations are technological, involving the pre- and post-translation stages, respectively. Both of them, functioning separately or together, produce no translation; they are contained within a single language. They are intralingual and non-machine operations as contrasted with the second bilingual A^1-B^1 or second multilingual $A^1-(B^1+C^1+\ldots+K^1)$ operations, i.e., the strictly translation operations, which are performed directly by the machine. The resultant text does not need to have all the native grammatical forms inserted and may be used just as it comes from the machine. So my machine is a real translation machine.

Hostile critics of the idea of monolingual translation are vexed by the fact that before my machine can translate, the text must first be provided with a logical analysis in order to have the form A^1 (or in simultaneous translation the form B^1, $C^1 \ldots K^1$). They would like even this step to be eliminated.

But any meaningful text, regardless of the language, always has the symbols of a logical analysis and represents the native form of a logical analysis. All I want in translating by my method is to replace this native system, i.e., the system characteristic of each individual language, with another system, a common system, my system, which is simpler than any native system. And that is all.

10. Description of the Parameters of the Stages in Constructing P. P. Smirnov-Troyanskiy's Translation Machine

Stages	Parameters			
	Design of machine	Capacity of machine	Alternation of sentences in translation	Degree of readiness of project
First	A model functioning automatically	Limited	Compulsory	Completed in May 1941 and proved satisfactory
Second	Test model of an automatic electrical system assembled from available instruments and apparatus	Limited	Free	Next stage -- everything ready for assembling the machine
Third	A powerful working unit, designed in accordance with the principles of up-to-date communications	Unlimited	Free	Theoretical research for this stage completed; project for actual

technology and operating automatically. The size is equal to the Moscow region automatic telephone system (ATS)	construction worked out

11. Examples Illustrating the Translation Processes Based on the Principles of the Monolingual Method

The following is an explanation of the symbols of logical analysis used in the examples given below. The symbols could also be given as numbers.

- as = a predicate in the form of a verb, present tense, indicative mood
- i = infinitive
- o = subject in the form of a noun, nominative case, singular number
- de = preposition to indicate the genitive case of all declinable words
- ojn = direct object in the form of a noun in the accusative case, plural number, without a preposition (o - noun, j - plural, n - accusative)

(Other symbols are explained in Part II, Point 22.)

Example 1. Translation from French into Russian (according to formula 10)

A	A^1	B^1	B
Le parti	le parti - o	partiya - o	Partiya
périt	périr - as	pogibat' - as	pogibayet
s'	si	yesli	yesli
il	il	on	ona
commence	commencer - as	nachinat' - as	nachinayet
à cacher	cacher - i	skryvat' - i	skryvat'
ses	son - ajn	svoy - ajn	svoi
erreurs	l'erreur - ojn	oshibka - ojn	oshibki

Example 2. Translation from German into Russian and French simultaneously (according to formula 12)

A	A^1	(B^1+C^1)		(B+C)	
Das Bild	Das Bild - o	kartina - o	le tableau - o	Kartina	Le tableau
der Welt	die Welt de - o	mir - de - o	le monde de - o	mira	du monde
zeigt,	zeigen - as	pokazyvat' - as	montrer - as	pokazyvayet,	montre,
wie	wie	kak	comment	kak	comment

die Materie	die Materie - o	materiya - o	la matière - o	materiya	la matière
sich bewegt,	sich bewegen - as	dvigat'sya - as,	se mouvoir - as	dvizhet-sya	se meut,
wie	wie	kak	comment,	kak	comment,
die Materie	die Materie - o	materiya - o	la matière - o	materiya	la matière
denkt.	denken - as.	myslit' - as.	penser - as.	myslit.	pense.

12. Use of Examples to Examine the Features of the Monolingual Method

These examples may serve to illustrate the way all 12 translation procedures in the monolingual method are constructed in accordance with Table 1. The texts are arranged in columns as follows: for example 1 -- (1) A, (2) A^1, (3) B^1, and (4) B; for example 2 -- (1) A, (2) A^1, (3) B^1+C^1, and (4) B+C.

Transposing the columns in various ways and taking them in twos, threes, and fours, in accordance with the monolingual procedures outlined in Table 1, we shall have in front of us materials to illustrate the functioning of the formulas in all four monolingual translation groups as well as in the fifth group.

Let us use the formulas from the comparative table.

The first group of procedures can be examined in the combinations of columns (each combination is to be taken from the same example): (1) A^1-B^1, (2) $B-A^1$, and (3) $A^1 - (B^1+C^1)$, i.e., according to formulas 1, 2, and 3 in the table.

The second group can be traced in the combinations of columns: (1) A^1-B^1-C, (2) B^1-A^1-A, and (3) $A-(B+C)-(B+C)$, i.e., according to formulas 4, 5, and 6.

The third group can be traced in the combinations: (1) A-A-B, (2) B-B-A, and (3) A-A-(B+C), i.e., according to formulas 7, 8, and 9.

The fourth group can be traced in the combinations: (1) A-A-B-B, (2) B-B-A-A, and (3) A-A-(B+C)-(B+C), i.e., according to formulas 10, 11, and 12.

The fifth group, including all three procedures of the bilingual method, can also be shown by combining the columns as follows: (1) A-B, (2) B-A, and (3) A-(B+C), i.e., according to formulas 13, 14, and 15.

Translation according to formula 1		Translation according to formula 2	
A^1	B^1	B^1	A^1
le parti - o	partiya - o	partiya - o	le parti - o
périr - as	pogibat' - as	pogibat' - as	périr - as
si	yesli	yesli	si
il	on	on	il
commencer - as	nachinat' - as	nachinat' - as	commencer - as
cacher - i	skryvat' - i	skryvat' - i	cacher - i

son - ajn	svoy - ajn	svoy - ajn son - ajn
l'erreur - ojn	oshibka - ojn	oshibka - ojn l'erreur - ojn

Translation according to formula 4

A^1	B^1	B
le parti - o	partiya - o	Partiya
périr - as	pogibat' - as	pogibayet
si	yesli	yesli
il	on	on
commencer - as	nachinat' - as	nachinayet
cacher - i	skryvat' - i	skryvat'
son - ajn	svoy - ajn	svoi
l'erreur - ajn	oshibka - ojn	oshibki

Translation according to formula 5

B^1	A^1	A
partiya - o	le parti - o	Le parti
pogibat' - as	périr - as	périt
yesli	si	s'
on	il	il
nachinat' - as	commencer - as	commence
skryvat' - i	cacher - i	à cacher
svoy - ajn	son - ajn	ses
oshibka - ojn	l'erreur - ojn	erreurs

Translation according to formula 7

A	A^1	B^1
Le parti	le parti - o	partiya - o
périt	périr - as	pogibat' - as
s'	si	yesli
il	il	on
commence	commencer	nachinat' - as
à cacher	cacher - i	skryvat' - i
ses	son - ajn	svoy - ajn
erreurs	l'erreur - ojn	oshibka - ojn

Translation according to formula 8

B	B^1	A^1
Partiya	partiya - o	le parti - o
pogibayet	pogibat' - as	périr - as

yesli	yesli	si
ona	on	il
nachinayet	nachinat^{9} - as	commencer - as
skryvat^{9}	skryvat^{9} - i	cacher - i
svoi	svoy - ajn	son - ajn
oshibki	oshibki - ojn	l^{e}erreur - ojn

Translation according to formula 11

B	B^1	A^1	A
Partiya	partiya - o	le parti - o	Le parti
pogibayet	pogibat^{9} - as	périr - as	périt
yesli	yesli	si	si
ona	on	il	il
nachinayet	nachinat^{9} - as	commencer - as	commence
skryvat^{9}	skryvat^{9} - i	cacher - i	à cacher
svoi	svoy - ajn	son - ajn	ses
oshibki	oshibka- ojn	l^{e}erreur - ojn	erreurs

13. Checks of Effectiveness of the Monolingual Translation Method Made Prior to Construction of the Machine (Shown in Charts)

1. Check in one's native language by each person himself.

Meaning of abbreviations in chart of check No 1: SLA-A = table of the symbols of logical analysis, with explanations in language A.

	At the first stage of checking	At the second stage of checking
1. Worker	T^A	T^A (the same)
2. Operation	$A\text{-}A^1$, regarded as the first of a total of three operations constituting the direct translation process, namely:	$A^1\text{-}A$, regarded as the third of a total of three operations constituting the reverse process, namely:
	$A\text{-}A^1$	$B\text{-}B^1$
	+ $A^1\text{-}B^1$	+ $B^1\text{-}A^1$
	$B^1\text{-}B$	$A^1\text{-}A$
	= $A\text{-}A^1\text{-}B^1\text{-}B$	= $B\text{-}B^1\text{-}A^1\text{-}A$
3. Tools and work objects	A SLA-A	A^1 SLA-A
4. Result	A^1	A

2. Check in two languages -- three persons engaged in the check. Meaning of abbreviations in chart of check No 2:

(1) The letter O = operator working on the model of a dictionary field. Cf. (3) below.

(2) SLA-A and SLA-B = tables of symbols of logical analysis, with explanations in languages A and B.

(3) DF (A+B) = stationary model of a dictionary field DF, i.e., a list of words in their basic forms in languages A and B at the same time. Cf. (1) above, on operator O.

(4) T with superscript A or B = monolingual translators working on language A or B, respectively.

(5) The function of operator O can be taken over by either the first or second monolingual translator. Only two participants in the translation process will then remain.

1.	Workers	T^A	O	T^B
2.	Operations	$A-A^1$	A^1-B^1	B^1-B
3.	Tools and work objects	A SLA-A	A^1 DF (A+B)	B^1 SLA-B
4.	Result	A^1	B^1	B

3. Check of translation from language A into several languages (A+B+C+...K = T) with T+1 monolingual translators, in addition to special operator O.

Meaning of abbreviations in chart for check No 3:

(1) O = operator working on the model of a dictionary field. Cf. (3) below.

(2) SLA-A, SLA-B, ... SLA-K = tables of symbols of logical analysis, with explanations in languages A, B, C, ...K.

(3) DF (A+B+C...+K) = stationary model of a dictionary field DF, i.e., a list of words in their basic forms in languages A, B, C, ..., K at the same time. Operator O works on this model of the dictionary field. Cf. (1) above.

(4) T with superscript A, B, C, ...,K = monolingual translators working on languages A, B, C,...K, respectively.

(5) The function of operator O can be taken over by any of the monolingual translators. The number of participants in the multilingual translation process will then correspond precisely to formula A+B+C+...+K = T+1 translators.

1.	Workers	T^A	O	T^B	T^C ...T^K
2.	Operations	$A-A^1$	$A^1-(B^1+C^1+...+K^1)$	B^1-B	$C^1-C...K^1-K$

3. Tools and work objects	A SLA-A	A^1 DF (A+B+C...+K)	B^1 SLA-B	C^1 ...K^1 SLA-C...SLA-K	
4. Result	A^1	$B^1+C^1...+K^1$	B	C...K	

14. How the Logical Analysis of a Native Text May Be Given in Footnotes, i.e., How to Combine a Native Text With a Logical Analysis (To Combine it Without Distortion of the Text)

NATIVE TEXT: Le[1] parti[1] périt[2] s'[3] il commence[4] à[5] cacher[5] ses[6] erreurs[7]. Le[1] tableau[1] du[8] monde[8] montre[4], comment la[1] matière[1] se[9] meut[9], comment la[1] matière[1] pense[4].

FOOTNOTES: [1] - o, [2] - ir - as, [3] si, [4] - er - as, [5] - er - i, [6] son - ajn, [7] - r - ojn, [8] de - o, [9] se mouvoir - as.

Note 1. Since the types of grammatical forms are repeated in a text and many identical cases are referred to the same footnotes, there cannot be too many footnotes on a page. Cf. footnote 1 as an example.

Note 2. There are three possible ways of using the symbols of logical analysis in footnotes.

Case 1. The form of the word in the text differs only slightly from its basic grammatical form. Each symbol of the logical analysis in the footnote is then preceded by one or two letters showing the original ending of the word. Cf. footnotes 2, 4, 5, 7.

Case 2. Only the basic grammatical form of the word is used in the text. The footnote then need contain only the symbol of logical analysis. Cf. the footnotes under 1.

Case 3. The textual form of the word differs completely from its basic grammatical form. In this case the latter is given in toto in the footnote before the symbol of logical analysis.

Note 3. The latest dictionaries present in regular alphabetical order all the basic forms of words coming from the same root or with the same meaning when the dissimilarity results from changes for number, degree of comparison, etc. They also include all the orthographical variants of irregular verbs. That is why one can find, for example, se meut referring to the basic form se mouvoir. This approach, in an improved version, will be widely used by the machine, thereby reducing the size and number of footnotes discussed here. Synonyms and idioms can be handled in similar fashion when a text is combined with a logical analysis.

APPENDIX

Recommendation of Professor I. D. Udal'tsov

The machine to produce translations in several languages at the same time is of great importance because of its underlying principle. I believe that there is much to be done in this connection and the results may well have an immense variety of applications. Comrade

Troyanskiy should be given financial support to enable him to construct his machine, if only in miniature.

[signed]
Ivan Udal'tsov
Rector, Moscow University

12 September 1955

Recommendation of Academician S. I. Vavilov

Academician
Sergey Ivanovich Vavilov

Moscow, B. Kaluzhskaya, 14

Tel. No.: B-1-41-90
30 October 1946

I am familiar with Comrade Troyanskiy's proposal for a translation machine. I consider it worthy of serious attention. If soundly executed, it may well produce a simple translation machine. It is desirable to construct a model for two or three languages and a dictionary of about 1,000 words.

[signed]
S. I. Vavilov

COMMENTARY

P. P. Smirnov-Troyanskiy's paper, "A Translation Machine Based on the Monolingual Method," is the fullest account we have of the linguistic principles underlying the construction of the earliest version of a translation machine. Written in February 1947, it contains a systematic description of all the ideas that he expressed as far back as 1933 and which were used as the starting point for the technical work from 1933 to 1939.

Nothing has been changed in or added to Troyanskiy's paper. The only omissions are some polemics directed at various persons with whom he had dealings.

The appendix contains two endorsements of Troyanskiy's work.

According to Professor W. Weaver, research on machine translation in the West started in 1946. Two years later A. D. Booth announced his intention of preparing a mechanical dictionary.[1] [Footnotes for this section are grouped together at the end of the section.]

Many of Troyanskiy's views later became the focus of attention of machine translation researchers. We should like to note the following:

(1) The Common Logical Structure of All Languages as a Basis of Machine Translation

Troyanskiy's paper begins with these words:

"In developing a new method of translation I started with the logical structure common to all languages.... This common logical structure brings languages close together, despite the variety of

structures, grammars, and vocabularies, and makes it possible to overcome the differences between them. On the basis of this unity I created a so-called form of logical analysis common to all languages as a kind of intermediate text in the translation process."[2]

It is interesting to compare this with the latest views of investigators of the same problem.

"A correlation of different languages is possible only if they share certain aspects. All languages actually do have a number of features in common .... Agreements in features which concern the logical aspect of language are especially numerous."[3]

"Now there are surely alogical elements in language (intuitive sense of style, emotional content, etc.)," Professor Weaver observes, "so that again one must be pessimistic about the problem of literary translation. But insofar as written language is an expression of logical character, this problem (i.e., machine translation -- I.B.) is at least formally solvable."[4]

To Weaver the most promising approach is one that "goes so deeply into the structure of languages as to come down to the level where they exhibit common traits."[5] This will establish that "universal language" which underlies all the extant living languages.

Weaver suggested that this "common language" could become an intermediate language and that translation might be effected in two stages: to "descend" from one language to the common base (the common language) and then "ascend" to some other language.[6]

Let us note in this connection that the idea of using the elements common to all languages that supposedly constitute some still undiscovered "common language" cannot, in our opinion, completely solve the problem of machine translation. It is the result of confusing the categories of language with the categories of logic.

In an attempt to isolate the invariant components of sentences in different languages, L. Wundheiler introduces the concept of "synonymous sentences" which he defines in such a way that the peculiarities of language connections and word relations in the sentence are reduced to zero.[7]

"Sentences are synonymous," Wundheiler writes, "insofar as they contain the same information about the relations between the objects (designata) mentioned therein. These relations are the common elements in synonymous sentences. Thus, "I gave a book to him" and "he was given a book by me" state the same relations between the person giving the book, the person to whom this book is given, and the book given. We believe that all languages have an element designating the action called in English "give" and that this element is used with some other elements designating the person giving, the person to whom something is given, and the thing given."

It is interesting to compare the examples of mechanical translation obtained by Troyanskiy's method and mentioned in his paper[8] with the examples of the "word-for-word translation" cited by Booth:[9]

(German)

Wenn in ein em gröss eren Gebiet zwei Form en neben einander leb en, ohne sich zu vermisch en, so gehor en sie verschieden en Form en kreis en an.

if in a/one d large (more) area two form n beside one another live z without self to/too mix z, so belong/hear p z z different m form m circle m at.

If, in a largish area, two forms coexist without intermixture, they will belong to different form-cycles.

It will be noted that besides a system of letter symbols of logical analysis Troyanskiy also suggested a number system in which each symbol has its own code number (cf. example in Section 14 above).

Finally, we should like to quote Oettinger on an automatic dictionary:

"An automatic dictionary is essentially a device performing the lexical translation only. The necessary grammatical translation must be performed either by other organs of a more complex machine or else by a person, preferably the ultimate reader himself."[10]

This substantially coincides with Troyanskiy's view.

(2) Editing Translations Produced by a Machine

Troyanskiy designed his machine translation system on the basis of "monolingual editors." Although it is now apparent that wholly automatic machine translation requiring no editing is quite feasible, researchers were long convinced that special editing was a prerequisite of machine translation.

Expressing the common view, E. Reifler wrote: "We believe that the machine, at least in the present stage of development of computer technique, cannot be placed between the foreign language text and the target language reader. This can be done only between the reader interpreting the foreign text in the same language and another reader to whom this text may come after it is translated."[11]

The discussions were limited only to the question of whether automatic translation requires two editors (one of whom is to prepare the source text in advance, the other to edit the resultant translation) or, with proper organization, only one editor. If the latter, which of the two is more essential?

Bar-Hillel believed that it was more important for an editor to process the text after it was translated[12], whereas Reifler thought he should do so before translation[13].

The necessity of additional scientific editing of a translated text to verify the meanings of the specialized terms was also discussed[14].

The specific tasks of editors were studied to determine "which functions we will assign to the machine and which ones we will leave to man[15]." There was no doubt that several functions would be performed by man even with machine translation, just as before.

"If we conceive of creating machines to translate without human intervention," Professor Reifler wrote, "machine translation will be either impossible or impractical."[16]

Arguing for the necessity of special editing, Reifler stated that languages have many instances of graphically identical categories with different meanings. The following, according to Reifler, have no graphic expression for differences in meaning:[17]

(a) polysemantic words;

(b) formation of grammatical meanings without changing the words, e.g., the plural of English sheep which is constructed without adding an ending (i.e., without changing the graphic appearance of the word);

(c) homonymic grammatical endings (e.g., the final d in the English decided may indicate either the past tense or past participle of the verb);

(d) differences in word order that are not made sufficiently clear in writing.

Machine translation researchers intended to rely on formal word changes in the sentence, by which they understood only changes in the spelling of individual words. However, they found in all languages a great many different cases of "graphic-semantic inexplicitness."

Regarding it as impossible for the machine to overcome this type of difficulty, they suggested the employment of a special editor to supply the source text with additional graphic signs to eliminate the lexical and grammatical confusions.

Under these conditions machine translation was bound to result in the replacement of translators with editors understanding only their native language, the "monolingual translators" of Troyanskiy. Although this represented but a partial solution of the problem, the practical necessity of accelerating the translation of the mass of material now published in all languages was so great that the idea was deemed worthy of attention.

The problem of special editing in the case of multilingual machine translation was particularly acute. It was observed that the degree of "graphic explicitness" of comparable categories varies from language to language. Therefore, in order to make it possible to have uniform editing of a text, Professor Reifler proposed[18] that "all grammatical meanings in all the source texts be arbitrarily regarded as given in zero form whether or not they were graphically expressed." These could then be raised to a standard level of "graphic explicitness" by means of some "universal system of graphic changes."

The following two methods were suggested as possible universal systems of editing. First, the earlier "universal system of auxiliary signals of machine translation" was based on a set of supralinear signs used in French, Greek, Hebrew, etc. For Chinese it was proposed that special graphic devices be used to indicate meaningful tones. These were to be based on the most complex (and, consequently, according to Reifler, the most complete) dialect, Cantonese, which has nine tones.

Reifler later proposed another editing system for machine translation that he called "the universal orthography of machine translation."[19] In this new system all words in a sentence are to receive

an indication of the part of speech to which they belong. The following changes are therefore introduced into the spelling of words in the source text: a capital letter is substituted either for the first letter of a word (in the case of nouns), second letter (in the case of verbs), or third letter (in the case of adjectives), etc. For example, hiS fiNe Qualities bEing perFectly uNderstood and aPpreciated in thOse Regions where hiS Lot wAs cAst (Charles Dickens).

Special marks are to indicate semantic features of words, e.g., whether they are agents or actions, whether they serve as modifiers of nouns or verbs, etc.

A comparison of these proposals with Troyanskiy's reveals that while he did not foresee all the possible difficulties, he conceived the basic idea which the later investigators adopted.

(3) Eventual Spread of Habit of Logical Analysis

The assumptions that were made in the course of research on machine translation echo Troyanskiy's belief that "in time people will become increasingly able to handle a text in the form of a logical analysis."[20]

"We might be able to turn over the difficulties of semantic interpretation to the consumers of machine translation," Reifler wrote later, "i.e., to those who want foreign texts translated by machine. This approach is closely linked to the problem of a special orthography. In its most general form the spelling of a text for machine translation means that people interested in machine translation will supply a center engaged in this work with the material in a special form better adapted in language and presentation to machine translation than when in the ordinary form. The special language may be either completely different from the ordinary language or a modification of it."[21]

In a paper presented at the 1952 conference on machine translation and published in Machine Translation of Languages, Stuart Dodd proposed a "Model English." The following are the most noteworthy among the ten rules that constitute, in his words, the "semantic ideal of 'one-word-one-meaning' or maximum one-to-one correspondence of a symbol and its referent":

(4) Every word should have just one order: i.e., one sequence in a sentence;

(5) Every word should have just one meaning -- an ideal seldom realized;

(6) Every word should have just one form: i.e., uninflected particles, or be for all tenses and persons of the copula;

(7) Every word should have just one pronunciation: the dominant one in ordinary speech;

(8) Every word should have just one spelling: preferably phonetic;

(9) Every letter should have just one sound: such as in the international phonetic alphabet.[22]

Quite naturally, many researchers linked the idea of "standardizing" a language for machine translation with the need of preliminary editing of the original.[23]

For a long time Reifler maintained that "determination of correct ungrammatical meaning might be assigned to the preliminary editor. He must be familiar with the language of the source text so that by adding marks such as diacritics he increases the semantic explicitness of the ordinary graphic form of the source text, thereby enabling the machine to select the correct equivalent in every case."[24]

Wundheiler suggested that a text should be edited "syntactically" by numbering the words:

I$_5$ know$_4$ now$_1$ well$_3$ he$_3$ said$_7$ yesterday$_6$ she$_{10}$ wished$_9$ poor$_{11}$ John$_{12}$ good$_{13}$ luck$_{14}$.

In "standardized" language:

Now very well know I yesterday said he wished she poor John good luck.

Proposals of this kind lead, of course, to a palliative solution of the problem. We can now say, with confidence, that machine translation, at least of scientific and technical literature, is feasible and does not require any artificial preparation of the language. However, what deserves the highest praise is the scientific foresight of P. P. Smirnov-Troyanskiy who anticipated by more than ten years the research that marked the beginning of extensive work on machine translation.

I. K. Bel'skaya

Footnotes

1. Machine Translation of Languages, New York, 1955.
2. Cf. the first paragraph of Troyanskiy's text, above.
3. E. Reifler, "The Mechanical Determination of Meaning" (MT of Languages, p. 140.
4. Warren Weaver, "Translation" (MT of Languages, p. 22).
5. Machine Translation of Languages, p. 23.
6. Ibid.
7. L. Wundheiler, "Invariant Prerequisite of All Translations" (MIT Library).
8. Cf. examples in Section 12, above.
9. Machine Translation of Languages, p. 37.
10. Ibid., p. 53.
11. E. Reifler, "Studies in Mechanical Translation," No 1 (MIT Library).

12. Bar-Hillel, "The Present State of Research in MT," (MIT Library).
13. E. Reifler, "Studies in Mechanical Translation," No 3 (MIT Library).
14. V. Oswald, "Microsemantics" (MIT Library); "Word-by-Word Translation," ibid.
15. Ibid.
16. E. Reifler, "Studies in MT," No 3 (MIT Library).
17. Ibid., No 1
18. E. Reifler, "Studies in Mechanical Translation," No 1, 84 (MIT Library).
19. Ibid, No 3.
20. Cf. Section 8 above.
21. E. Reifler, "The Mechanical Determination of Meaning" (MT of Languages, p. 142). Cf. also Wundheiler's remarks on the possibility of creating a "normal language" with this mandatory syntactic rule: "The complements of any expression whatsoever must follow in the same fixed order" (Wundheiler, "Invariant Prerequisite of All Translations," MIT Library). For example "give" must be immediately followed by the giver, then by the one to whom something is given, then by the object given.
22. Stuard Dodd, "Model English," (MT of Languages, p. 166).
23. Editing at the early stages of automatic translation is an important independent problem.
24. E. Reifler, "The Mechanical Determination of Meaning" (MT of Languages, p. 141).
25. L. Wundheiler, "Invariant Prerequisite of all Translations (MIT Library).

PART II. TECHNICAL MATERIALS

P. TROYANSKIY

A MACHINE TO PRODUCE AUTOMATICALLY FINISHED PRINTED TRANSLATIONS FROM ONE INTO SEVERAL LANGUAGES SIMULTANEOUSLY AND REQUIRING ONLY LITERARY EDITING

29 July to 3 September 1933

Moscow

1. A frame 1 tilted at an angle of 60° is supported by end stands. Its design is such that a belt can go around it in two directions. (1) rectilinearly to the right and to the left, and (2) vertically (around the frame). All the bearings are of the roller and ball types. The surface of the belt is flat with a number of openings 3 like the finger holes on the dial of an automatic telephone.

2. Thus, any point on the surface of the belt can be set opposite aperture 6 of camera 5.

3. Pasted to the surface of the belt is, let us say, a six- (or more) language parallel dictionary in alphabetical order lying flat and unfolded. The columns of words are so arranged that words with the most frequently used letters (k, m, etc.) are close to the center (the same principle as that of the typewriter keyboard or type case). The dictionary must be specially prepared for this machine (cf. point 12 below). The belt surface with the attached dictionary is called the "vocabulary field."

4. The table with the camera standing in front of the vocabulary field also has a typewriter containing two types (e.g., Russian and Latin) and a set of special keys called the "keys of logical and etymological analysis" (see point 22 below).

5. A tape consisting of two strips — the left one with an emulsified surface, the right one with ordinary typewriting paper, or vice versa — passes through the typewriter and camera. Both apparatuses operate on one line, each in its own portion. Further, if the paper or emulsified parts of the tape are pasted together, the typewriter must be so constructed that the paper does not run off the machine to the left, but the machine is to travel to the right by itself (perhaps together with the operator's chair!). It may be possible to design things so that both parts of the tape can be pasted together later. In this case their movements will have to be synchronized. However, there will be no need of a special typewriter.

6. The actual work is performed in this way. Shifting the vocabulary field upwards and sideways by the finger holes on the belt surface (the field moves as easily as the free wheel of a chain bicycle), the operator sets the desired word opposite aperture 6 of the camera (view finder) and halts the dictionary field at that place by an electromagnetic brake with a "stop" button. He then releases the shutter, exposes the lines, and prints the symbols of logical analysis on the paper part of the tape. After a special key is pressed, the tape (both the emulsified and paper parts) is moved one line upward, freeing the place for the next writing and exposure.

7. The vocabulary field belt is then released and the operation starts all over again.

8. If electrical rather than manual power is used, all the operations mentioned above are carried out by press button control. Pressing a key of the special keyboard for the first letter of the word to be translated causes the belt automatically to move the required column along the line of aperture 6 in front of the camera. With pressure from the second letter of the word to be translated the belt is automatically set by the first word which starts with these two letters opposite the aperture. A third key moves the belt under the aperture by the first word which starts with this combination of three letters, etc. Pressing a special key causes the belt to turn slowly

until it selects the appropriate line, which is finally held by the "stop" key so that it can be photographed and the logical symbols inserted. Before the symbols are printed and the line photographed, the selected line must be automatically shown in enlarged form through the lens so that it can be checked. A special key then causes the exposure to be made and the line above translated. The apparatus is now ready for the next word. For the subsequent development of the machine, see points 28-31 [not given in this brochure].

9. The size of the vocabulary field: with a line for six languages -- 10 centimeters long; in height -- four short lines per centimeter. Twenty square meters can accommodate 80,000 roots. With these dimensions the surface of the face side of the belt equals 2 x 4 meters.

10. The following facts will help us to visualize what these 80,000 roots mean. L. Tolstoy used 12,000 roots in all his works. An educated man uses 2,000 in conversation. The Russian language (including all the historic and obsolete forms) has 180,000. English (including all those used at any time) has 200,000. Technical journals use 3,000.

11. To facilitate the search for a required line in the vocabulary field, a strong magnifying lens is mounted before the aperture and automatically deflected when a picture is taken.

12. How is the dictionary to be prepared before it is entered in the vocabulary field?

(1) Synonyms are written out in a line, for example:

speak	govorit', razgovarivat'	parler, causer	sprechen
swift	bystryy, skoryy	rapide, vite, soudain	schnell

13. The order in which the immediate output of the machine is processed.

The text comes from the machine in the form of a series of columns; the column on the extreme right, for example, is the logical analysis column, the others (depending on the number of languages) contain the basic lexical forms produced by the camera from the vocabulary field.

14. Further processing includes:

(2) Homonyms are given with an indication of the meaning in parentheses, for example:

Kosa (sand bar)	Perevod (transfer to another assignment)
Kosa (girl's braid)	Perevod (translation)
Kosa (scythe)	Perevod (money order)
	Perevod (shunting)
	Perevod (transfer of picture)

Note: Homonyms are found only in the dictionary of the language from which the translation is to be made.

(1) What does the typist do with the text as it comes from the machine? She copies the column of each language just as it comes from the machine and types alongside it a column of logical analysis symbols, being careful, of course, to match the lines correctly.

(2) What does the copy editor then do with the text? He combines the two columns -- that of lexical selections and that of logical analysis symbols -- into a single connected column of text.

(3) What does the literary editor do with this text? He checks the copy editor's work, crosses out the synonyms not needed, and polishes the translation.

15. What in essence is translation? Into what elements is the process broken down? These elements are as follows:

(1) Selection of the meanings of the roots of variable words and selection of the meanings of invariable words;

(2) Establishment of the logical relationship between the words, i.e., production of syntactic analysis, determination of the syntactic relationship (subject, predicate, attributive, object, adverbial modifiers);

(3) Establishment of intersubordination of forms between variable and invariable words (clauses).

16. The machine described: the first thing it does automatically; the third thing it does with the help of an operator; the second thing constitutes the operator's plan.

17. What languages do an operator (of the machine), copy editor, and editor need to know?

18. An operator must know only one language, the original, plus the system of logical symbols plus a small vocabulary of two to three hundred "auxiliary" words (in Esperanto, which, in this respect, is the most efficient and shortest; it is constructed on scientific principles and is not a spontaneous creation).

19. The copy editor (who combines the columns of words and logical symbols) must know only his native language. He leaves the synonyms for the literary editor, supplying only their grammatical form (case, number, gender, tense, mood, voice, etc.). Moreover, he too has to know the same set of "auxiliary" words (where, by the way, it will be useful to group all the Esperanto suffixes and prefixes to make the meaning more precise).

20. The literary editor will have to know the two languages involved in order to be able to check the translation, choose the synonyms, smooth out the rough spots, in short, to produce a finished piece of work.

21. Thus, the most laborious part of the translation is carried out by operators and copy editors who do not need to know even two languages! This has unusually important implications for translation from and into the languages of the minority nationalities of the Soviet Union.

22. The following are written on paper tape in the logical analysis column (with everything that can be written in the words from the dictionary of "auxiliary" words in Esperanto, i.e., from the 200 to 300 referred to above):

(1) numerals; (2) designations of concrete numbers (e.g., cu. m., kw., mlm., km., etc.); (3) pronouns; (4) conjunctions (with a few exceptions); (5) prepositions (with a few exceptions); (6) proper nouns (names of cities, rivers, seas, surnames, given names, etc.); (7) abbreviations and code words (NKTP, USSR, USA, LZ127, etc.); (8) so-called international terms, e.g., revolution, Marxist, radio, aluminum, etc., which differ orthographically from language to language; (9) symbols of logical analysis and inflection:

j - plural;
n - relation of form of inflected part of speech to verb (government of verb); (direct object); (accusative case);
de - (1) relation of form of inflected part of speech to another inflected word; (2) agent with passive voice;
per - instrument with passive voice (with inflected part of speech); (genitive case);
e - (1) adverb; (2) gerund;
oni - impersonal form of verb (German man, French on);
a - (1) meaning of an adjective with a predicate expressed by some form of an auxiliary verb; (2) participle;
i - infinitive;
as)
is) - past, present, and future tenses of verbs in the indicative
os) mood;
us - conditional mood;
u - imperative mood;
pli ol - comparative degree;
la plej el - superlative degree;
on)
obl) - fractions, multipliers, connectives, disjunctives;
op)
po)

ant)
int) - active participles, present, future, and past tenses;
ont)

at)
it) - passive participles, present, future, and past tenses.
ot)

(10) all words and expressions that the operator finds difficult to translate even descriptively, i.e., by the use of synonyms, are to be followed by an exclamation point or question mark in parentheses; (11) certain questions in parentheses about a given word which determine the

logical place of the word in a sentence; these questions are put in Esperanto, e.g., kies, kie, kien, alkiu, kial, kiam, kiom da, kioma [whose, where, whither, to whom, why, when, how much, which], etc; (12) all punctuation marks (comma, period, dash, parentheses, quotation marks, etc.); (13) invariable words generally require no logical analysis symbols; (14) some forms of invariable words also remain without symbols such as the nominative case, singular number, and infinitive.

Note: The time of processing and length of tape coming from the machine may be further shortened if a word is given no symbol and the following word does not have to be photographed. In this case the lines can be combined with a "+" sign.

Class 15.20 No 40995

57 1

CERTIFICATE FOR INVENTION

Description
of a Machine to Select and Print Words in Translating
From One Language Into Another

In connection with P. P. Troyanskiy's patent,
5 September 1933
(Priority No. 134430).
Published 31 January 1935.

The proposed machine is designed to select and print words in translating from one language into another. It consists chiefly of a special tape containing words in various languages that moves along a table. Openings in the tape are used to halt it in front of a camera which stands next to a typewriter provided with extra keys to print conventional symbols opposite the photographed word (see Figure 1).

As is evident from the sketch, the machine consists of a flat, inclined table 1 along which tape 2 can move freely in various directions. The tape has finger openings 3 by which it can be set opposite aperture 6. On the tape is pasted a six-language parallel dictionary arranged alphabetically. The dictionary lies flat and unfolded with the columns so arranged that the words with the most frequently used letters (k, m, p, etc.) are close to the center.

The machine works as follows. The tape is moved in such a way that the word corresponding to the word to be translated is set exactly opposite aperture 6, whereupon the tape is stopped and the shutter of the camera released for the word to be photographed on photosensitive tape drawn through the camera. The conventional symbols of logical analysis are simultaneously printed on paper tape. The tapes of the camera and typewriter are then moved one line and tape 2 shifted to handle the next word, and so on.

The typist puts together a connected text from the translation obtained in the form of two joined (or pasted together) tapes with columns of photographed words and conventional symbols of logical analysis. The text then goes to a copy editor who attaches expressions to the words corresponding to the symbols of logical analysis. After recopying it is turned over the literary editor for final editing.

The Invention

The machine to select and print words in translating from one language into another features a circular tape 2 with words in different languages pasted on it. The openings 3 are used to set the required word opposite an aperture in a table over which a camera is placed to photograph the main words on photosensitive tape. Alongside the camera is a typewriter equipped with extra keys to print on paper tape conventional symbols opposite the photographed words.

Expert and editor, A. G. Bremzen

FIGURE APPENDIX

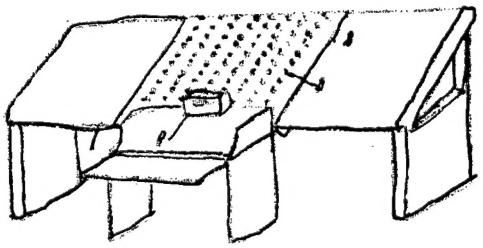

Figure 1

5214

END